

# Exact quantum query complexity of $\text{EXACT}_{k,l}^n$

Andris Ambainis<sup>1</sup>, Jānis Iraids<sup>1</sup>, and Daniel Nagaj<sup>2</sup>

<sup>1</sup>Faculty of Computing, University of Latvia, Raiņa bulvāris 19, Riga, LV-1586, Latvia,  
ambainis@lu.lv, janis.iraids@gmail.com

<sup>2</sup>Institute of Physics, Slovak Academy of Sciences, Dúbravská cesta 9, 845 11 Bratislava,  
Slovakia, dnagaj@gmail.com

January 26, 2017

## Abstract

In the decision tree model, one's task is to compute a boolean function  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  on an input  $x \in \{0, 1\}^n$  that is accessible via queries to a black box (the black box hides the bits  $x_i$ ). In the quantum case, classical queries and computation are replaced by unitary transformations. A quantum algorithm is exact if it always outputs the correct value of  $f$  (in contrast to the standard model of quantum algorithms where the algorithm is allowed to be incorrect with a small probability). The minimum number of queries for an exact quantum algorithm computing the function  $f$  is denoted by  $Q_E(f)$ .

We consider the following  $n$  bit function with  $0 \leq k \leq l \leq n$ :

$$\text{EXACT}_{k,l}^n(x) = \begin{cases} 1, & \text{if } x_1 + \dots + x_n \in \{k, l\}, \\ 0, & \text{otherwise,} \end{cases}$$

i.e. we want to give the answer 1 only when exactly  $k$  or  $l$  of the bits  $x_i$  are 1. We construct a quantum query algorithm for this function and give lower bounds for it, with lower bounds matching the complexity of the algorithm in some cases (and almost matching it in other cases):

- If  $l - k = 1$  and  $k = n - l$ , then  $Q_E(\text{EXACT}_{k,k+1}^{2k+1}) = k + 1$ .
- If  $l - k \in \{2, 3\}$ , then  $Q_E(\text{EXACT}_{k,l}^n) = \max\{n - k, l\} - 1$ .
- For all  $k, l$ :  $\max\{n - k, l\} - 1 \leq Q_E(\text{EXACT}_{k,l}^n) \leq \max\{n - k, l\} + 1$ .

## 1 Introduction

In this paper we study the computational complexity of boolean functions in the quantum black box model. It is a generalization of the decision tree model, where we are computing an  $n$ -bit function  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  on an input  $x \in \{0, 1\}^n$  that can only be accessed through a black box by querying some bit  $x_i$  of the input. In the quantum black box model the state of the computation is described by a quantum state from the Hilbert space  $\mathcal{H}_Q \otimes \mathcal{H}_W \otimes \mathcal{H}_O$  where  $\mathcal{H}_Q = \{|0\rangle, |1\rangle, \dots, |n\rangle\}$  is the query subspace,  $\mathcal{H}_W$  is the working memory and  $\mathcal{H}_O = \{|0\rangle, |1\rangle\}$  is the output subspace. A computation using  $t$  queries consists of a sequence of unitary transformations  $U_t \cdot O_x \cdot U_{t-1} \cdot O_x \cdot \dots \cdot O_x \cdot U_0$  followed by a measurement, where the  $U_i$ 's are independent of the input and  $O_x = O_{Q,x} \otimes I \otimes I$  with

$$O_{Q,x} |i\rangle = \begin{cases} (-1)^{x_i} |i\rangle = \hat{x}_i |i\rangle, & \text{if } i \in [n], \\ |0\rangle, & \text{if } i = 0, \end{cases}$$

is the query transformation, where  $x_i \in \{0, 1\}$  or equivalently,  $\hat{x}_i \in \{-1, 1\}$ . The final measurement is a complete projective measurement in the computational basis and the output of the algorithm is the result of the last register,  $\mathcal{H}_O$ . We say that a quantum algorithm computes  $f$  exactly if for all inputs  $x$

the output of the algorithm always equals  $f(x)$ . Let us denote by  $Q_E(f)$  the minimum number of queries over all quantum algorithms that compute  $f$  exactly.

For quite a long time the largest known separation between the classical decision tree complexity  $D(f)$  and  $Q_E(f)$  was only by a factor of two — the XOR of two bits can be computed exactly using only 1 quantum query [7, 8, 9]. However, in 2012 Ambainis gave the first asymptotic separation that achieved  $Q_E(f) = O(D(f)^{0.8675})$  for a class of functions  $f$  [1]. Next, in 2015 Ambainis et al. used pointer functions to show a near quadratic separation between these two measures:  $Q_E(f) = \tilde{O}(\sqrt{D(f)})$  [2]. On the other hand Midrijānis has proved that the maximum possible separation between  $Q_E(f)$  and  $D(f)$  is at most cubic [12].

However, the techniques for designing exact quantum algorithms are rudimentary compared to the bounded error setting. Other than the well known XOR *trick* — constructing a quantum algorithm from a classical decision tree that is allowed to “query” the XOR of any two bits — there are few alternate approaches. In addition to the asymptotic separations of [2, 1], Montanaro et al. [13] gave a 2 query quantum algorithm for the symmetric 4 bit function

$$\text{EXACT}_2^4(x) = \begin{cases} 1, & \text{if } x_1 + x_2 + x_3 + x_4 = 2, \\ 0, & \text{otherwise,} \end{cases}$$

and showed that it could not be computed optimally using the XOR trick. Afterwards Ambainis et al. gave an algorithm [3] for two classes of symmetric functions:

$$\text{EXACT}_k^n(x) = \begin{cases} 1, & \text{if } x_1 + x_2 + \dots + x_n = k, \\ 0, & \text{otherwise} \end{cases} \quad ; \quad Q_E(\text{EXACT}_k^n) \leq \max\{k, n - k\},$$

and the threshold function

$$\text{TH}_k^n(x) = \begin{cases} 1, & \text{if } x_1 + x_2 + \dots + x_n \geq k, \\ 0, & \text{otherwise} \end{cases} \quad ; \quad Q_E(\text{TH}_k^n) \leq \max\{k, n - k + 1\}.$$

For partial functions quantum algorithms with superpolynomial speedup are known [8, 6]. It seems that our work relates well to the results of Qiu and Zheng on partial functions based on the Deutsch-Jozsa problem [14].

## 1.1 Our results

We consider exact quantum algorithms for symmetric total boolean functions, i.e., functions for which permuting the input bits does not change its value. For symmetric functions, the largest known separation remains a factor of 2. We know from von zur Gathen’s and Roche’s work on polynomials [10] and quantum lower bounds using polynomials [4] that for symmetric  $f$  :  $Q_E(f) \geq \frac{n}{2} - O(n^{0.548})$ , thus the largest known separation is either optimal or close to being optimal.

However, many of the known exact algorithms are for symmetric functions (for example, XOR, EXACT and TH functions mentioned in the previous section). Because of that, we think that symmetric functions may be an interesting ground to explore new methods for developing more exact quantum algorithms.

In Section 3.1 we present an algorithm achieving up to  $D(f) = 2Q_E(f)$  for a certain class of symmetric functions

**Definition 1.** Let  $\text{EXACT}_{k,l}^n$  for  $0 \leq k \leq l \leq n$ , be an  $n$ -argument symmetric boolean function that returns 1 if and only if the input contains exactly  $k$  ones or exactly  $l$  ones.

$$\text{EXACT}_{k,l}^n(x) = \begin{cases} 1, & \text{if } |x| \in \{k, l\}; \\ 0, & \text{otherwise.} \end{cases}$$

Let us denote by  $d$  the separation between  $l$  and  $k$ :  $d = l - k$ . In general a symmetric boolean function  $\text{SYM}_a$  on  $n$  input bits can be defined by a list  $a = (a_0, \dots, a_n) \in \{0, 1\}^{n+1}$  such that  $\text{SYM}_a(x) = a_{|x|}$ . When  $d > 0$  it may be convenient to think of  $\text{EXACT}_{k,l}^n$  in this way. In this representation  $\text{EXACT}_{k,l}^n$

corresponds to lists  $a$  of length  $n + 1$  with two 1s and the number of zeroes before the first, after the last 1, and distance between 1s correspond to parameters  $k$ ,  $n - l$ , and  $d$  respectively.

The boundary cases,  $d = 0$  and  $d = n$ , have been solved previously. When  $d = n$ , the function is usually referred to as EQUALITY $_n$ . It can be solved with  $n - 1$  quantum queries using the well-known XOR trick. The case  $d = 0$  is also known as the EXACT $_k^n$  function which has been analyzed in [3] where it was shown that  $Q_E(\text{EXACT}_k^n) = \max\{k, n - k\}$ . In this paper, we completely solve the  $d \in \{2, 3\}$  cases and partially solve the  $d = 1$  case and  $d \geq 4$  case.

The first of our results is

**Theorem 1.** *If  $d = 1$ ,  $l = n - k$  and  $k > 0$ , then for  $\text{EXACT}_{k,l}^n = \text{EXACT}_{k,k+1}^{2k+1}$*

$$Q_E(\text{EXACT}_{k,k+1}^{2k+1}) = k + 1.$$

The algorithm we provide in the proof works also when  $l \neq n - k$  by padding the function. However, the algorithm is then only an upper bound on  $Q_E(\text{EXACT}_{k,k+1}^n)$ . For example,  $Q_E(\text{EXACT}_{2,3}^3) = 2$  but our algorithm uses 3 queries for the padded version of the function (if we pad the input with two zeroes, we end up computing  $\text{EXACT}_{2,3}^5$ ). Furthermore, the computations by Montanaro et al. [13] suggest that  $Q_E(\text{EXACT}_{3,4}^5) = 3$  and  $Q_E(\text{EXACT}_{4,5}^6) = 4$ . There, unlike the  $\text{EXACT}_{2,3}^3$  case, we don't know what the optimal algorithm looks like.

Next, we have a complete understanding of the  $d \in \{2, 3\}$  case,

**Theorem 2.** *If  $d \in \{2, 3\}$ , then*

$$Q_E(\text{EXACT}_{k,l}^n) = \max\{n - k, l\} - 1.$$

In particular, when  $d = 2$  and  $l = n - k$ , we have  $l = k + 2$  and  $n = 2k + 2$ , meaning  $l = \frac{n}{2} + 1$ , giving us  $Q_E(\text{EXACT}_{k,l}^n) = \frac{n}{2}$  whereas the deterministic query complexity is  $D(\text{EXACT}_{k,l}^n) = n$ , hence we exhibit a factor of 2 gap between  $Q_E(f)$  and  $D(f)$  which is the largest known gap for a symmetric boolean function.

For larger values of  $d$ , we provide an exact quantum algorithm and a lower bound that is 2 queries less than the complexity of the algorithm:

**Theorem 3.** *If  $d \geq 4$ , then*

$$\max\{n - k, l\} - 1 \leq Q_E(\text{EXACT}_{k,l}^n) \leq \max\{n - k, l\} + 1,$$

We conjecture that our lower bound is tight, i.e., that

**Conjecture 1.** *If  $d \geq 2$ , then*

$$Q_E(\text{EXACT}_{k,l}^n) = \max\{n - k, l\} - 1.$$

The lower bound of Theorem 3 combined with Theorem 1 implies that

$$Q_E(\text{EXACT}_{k,l}^n) \geq \frac{n}{2}.$$

Interestingly, the algorithm of Theorem 3 can be used to compute a wide variety of symmetric functions with asymptotically optimal number of queries. Namely, we show

**Theorem 4.** *Let  $a \in \{0, 1\}^{n+1}$  be a binary string with no 1-s far from its center, i.e. there exists some  $g(n) \in o(n)$  such that  $|i - \frac{n}{2}| > g(n) \implies a_i = 0$ . Then,*

$$Q_E(\text{SYM}_a) = \frac{n}{2} + o(n).$$

Since  $D(\text{SYM}_a) = n$  for any such function  $\text{SYM}_a$  (except for one that is 0 on all inputs), we obtain a factor- $(2 - o(1))$  advantage for exact quantum algorithms for any such  $\text{SYM}_a$ .

The outline for the rest of the paper is as follows. We describe the lower bound parts of Theorems 1, 2 and 3 in section 2 and the algorithms for these theorems in section 3. The algorithm for Theorem 4 is given in Appendix B.

## 2 The lower bounds

### 2.1 Proofs of the lower bound theorems

**Theorem 5.** *If  $d \geq 1$ , then*

$$Q_E(\text{EXACT}_{k,l}^n) \geq \max\{n - k, l\} - 1.$$

This theorem provides the lower bound part for Theorems 2 and 3.

*Proof of Theorem 5.* Consider the function  $\text{EXACT}_{k,l}^n$  with  $l \leq n - k$  ( $l \geq n - k$  is symmetric and gives the  $l - 1$  result in the theorem). If the first  $k$  input bits are ones, a quantum algorithm computing  $\text{EXACT}_{k,l}^n$  must be computing  $\text{EXACT}_{0,l-k}^{n-k}$  on the remaining  $n - k$  input bits. Next we proceed similarly as in the lower bound via polynomials for  $\text{OR}_n$  function[4]. There must exist a state  $|\psi(x)\rangle \in \mathcal{H}_Q \otimes \mathcal{H}_W \otimes |1\rangle$  which for  $x = (0, \dots, 0)$  is non-zero at the end of the computation. If the algorithm performs  $t$  queries, then the amplitude of the state  $|\psi(x)\rangle$  can be expressed as a degree  $\leq t$  multilinear polynomial in  $\hat{x}$ :

$$p(\hat{x}_1, \dots, \hat{x}_n) = \sum_{\substack{S: S \subseteq [n] \\ |S| \leq t}} \alpha_S \prod_{i \in S} \hat{x}_i.$$

Let  $p_{\text{sym}}$  be the symmetric polynomial

$$p_{\text{sym}}(\hat{x}_1, \dots, \hat{x}_n) = \sum_{\pi \in S_n} \frac{p(\hat{x}_{\pi(1)}, \dots, \hat{x}_{\pi(n)})}{n!}.$$

Crucially, for the inputs  $x \in \{(0, \dots, 0)\} \cup \{x | \text{EXACT}_{0,l-k}^{n-k}(x) = 0\}$ :

$$p_{\text{sym}}(\hat{x}_1, \dots, \hat{x}_n) = p(\hat{x}_1, \dots, \hat{x}_n).$$

By assigning  $s := \frac{n - (\hat{x}_1 + \dots + \hat{x}_n)}{2}$  we can obtain a polynomial  $q(s)$  that for all  $\hat{x} \in \{-1, 1\}^n$ :

$$q\left(\frac{n - (\hat{x}_1 + \dots + \hat{x}_n)}{2}\right) = p_{\text{sym}}(\hat{x}_1, \dots, \hat{x}_n).$$

The polynomial  $q$  is therefore non-zero on  $s = 0$  and zero on  $s \in \{0, 1, \dots, n - k\} \setminus \{0, l - k\}$ . Thus it is a non-zero polynomial of degree at least  $n - k - 1$ . On the other hand the degree of  $q$  is at most  $t$ . Thus  $n - k - 1 \leq \deg q \leq t$ .  $\square$

This lower bound is not tight when  $d = 1$  and  $l = n - k$ . In this case we use a more sophisticated approach and give a different but possibly more insightful proof.

**Theorem 6.** *If  $d = 1$ ,  $n > 1$  and  $l = n - k$ , then for  $\text{EXACT}_{k,l}^n = \text{EXACT}_{k,k+1}^{2k+1}$*

$$Q_E(\text{EXACT}_{k,k+1}^{2k+1}) \geq k + 1.$$

Theorem 6 yields a lower bound that is better by one query than Theorem 5, which yields a lower bound of  $k$ .

To show Theorem 6, we use an unpublished result by Blekherman.

**Theorem 7** (Blekherman). *Let  $q(\hat{x})$  be the symmetrization of a polynomial  $p^2(\hat{x}_1, \dots, \hat{x}_n)$  where  $p(\hat{x})$  is a polynomial of degree  $t \leq \frac{n}{2}$ . Then, over the Boolean hypercube  $\hat{x} \in \{-1, 1\}^n$ ,*

$$q(\hat{x}) = \sum_{j=0}^t p_{t-j}(|x|) \left( \prod_{0 \leq i < j} (|x| - i)(n - |x| - i) \right)$$

where  $p_{t-j}$  is a univariate polynomial that is a sum of squares of polynomials of degree at most  $t - j$  and  $|x|$  denotes the number of variables  $i : \hat{x}_i = -1$ .

See [11] for a proof of Blekherman's theorem. Furthermore, we provide a considerably shorter proof in the next subsection.

*Proof of Theorem 6.* Let us consider the negation of the function  $\text{EXACT}_{k,k+1}^{2k+1}$ . Assuming, towards a contradiction, that there exists a quantum algorithm computing the function with  $k$  queries, there exists a sum of squares representation of  $\text{NOT-EXACT}_{k,k+1}^{2k+1}$ :

$$\text{NOT-EXACT}_{k,k+1}^{2k+1}(x) = \sum_i r_i^2(\hat{x}),$$

such that  $\deg r_i \leq k$ . Since the function is symmetric, the symmetrization is also a valid representation. Since  $\text{Sym}(\sum_i r_i^2(\hat{x})) = \sum_i \text{Sym}(r_i^2(\hat{x}))$ , it follows from Blekherman's theorem that there is a univariate polynomial of the form

$$q(|x|) = \sum_{j=0}^k p_{k-j}(|x|) \left( \prod_{i=0}^{j-1} (|x|-i)(n-|x|-i) \right), \quad (1)$$

where  $q(|x|) = \text{NOT-EXACT}_{k,k+1}^{2k+1}(x)$  on the Boolean hypercube and  $p_{k-j}$  are sum of squares polynomials with  $\deg p_{k-j} \leq 2k - 2j$ . The polynomial  $q(|x|)$  is non-negative in the interval  $|x| \in [k-1, k+2]$ . Since the polynomial is 0 at  $|x|=k$  and  $|x|=k+1$ , it must have at least 3 local extrema in the interval  $|x| \in [k, k+1]$ . Additionally, it is 1 when  $|x| \in \{0, 1, \dots, n\} \setminus \{k, k+1\}$ , hence it has  $2k-2$  more extrema. In total the polynomial has at least  $2k+1$  local extrema, therefore its degree is at least  $2k+2$ . On the other hand by our assumption  $\deg q \leq 2k$  which is a contradiction.  $\square$

## 2.2 Proof of Blekherman's theorem

### 2.2.1 Group representation

Let  $H_\varphi$  be a Hilbert space with basis states  $\hat{x}_S$  (for all  $S \subseteq [n]$ ) corresponding to monomials  $\prod_{i \in S} \hat{x}_i$ . Then, the vectors in  $H_\varphi$  correspond to multilinear polynomials in variables  $\hat{x}_i$ . We consider a group representation of the symmetric group  $\mathfrak{S}_n$  on  $H_\varphi$  with transformations  $U_\pi$  defined by  $U_\pi \hat{x}_S = \hat{x}_{\pi(S)}$ . The irreducible representations contained in  $H_\varphi$  are well known:

Let  $S_m(\hat{x}_1, \dots, \hat{x}_n) = \sum_{i_1, \dots, i_m} \hat{x}_{i_1} \dots \hat{x}_{i_m}$  be the  $m^{\text{th}}$  elementary symmetric polynomial. We use  $S_0(\hat{x}_1, \dots, \hat{x}_n)$  to denote the constant 1.

**Lemma 1.** *A subspace  $H \subseteq H_\varphi$  is irreducible if and only if there exist  $b$  and  $\alpha_m$  for  $m = 0, 1, \dots, n-2b$  such that  $H$  is spanned by vectors  $\vec{p}_{i_1, \dots, j_b}$  corresponding to polynomials  $p_{i_1, \dots, j_b}$  (for all choices of pairwise distinct  $i_1, j_1, \dots, i_b, j_b \in [n]$ ) where*

$$p_{i_1, \dots, j_b}(\hat{x}_1, \dots, \hat{x}_n) = (\hat{x}_{i_1} - \hat{x}_{j_1}) \dots (\hat{x}_{i_b} - \hat{x}_{j_b}) \sum_{m=0}^{n-2b} \alpha_m S_m(\hat{x}')$$

and  $\hat{x}' \in \{-1, 1\}^{n-2b}$  consists of all  $\hat{x}_i$  for  $i \in [n]$ ,  $i \notin \{i_1, \dots, j_b\}$ .

See [5] for a short proof of Lemma 1.

### 2.2.2 Decomposition of $q(\hat{x})$

Let

$$p(\hat{x}_1, \dots, \hat{x}_n) = \sum_{S: |S| \leq t} a_S \hat{x}_S.$$

We associate  $p^2(\hat{x}_1, \dots, \hat{x}_n)$  with the matrix  $(P_{S_1, S_2})$  with rows and columns indexed by  $S \subseteq [n]$ ,  $|S| \leq t$  defined by  $P_{S_1, S_2} = a_{S_1} a_{S_2}$ . Let  $\vec{x}$  be a column vector consisting of all  $\hat{x}_S$  for  $S: |S| \leq t$ . Then,  $p^2(\hat{x}_1, \dots, \hat{x}_n) = \vec{x}^T P \vec{x}$ . This means that  $P$  is positive semidefinite.

For a permutation  $\pi \in \mathfrak{S}_n$ , let  $P^\pi$  be the matrix defined by

$$P_{S_1, S_2}^\pi = a_{\pi(S_1)} a_{\pi(S_2)}$$

and let  $Q = \frac{1}{n!} \sum_{\pi \in \mathfrak{S}_n} P^\pi$  be the average of all  $P^\pi$ . Then,  $q(\hat{x}) = \vec{x}^T Q \vec{x}$ .  $Q$  is also positive semidefinite (as a linear combination of positive semidefinite matrices  $P^\pi$  with positive coefficients).

We decompose  $Q = \sum_i \lambda_i Q_i$  with  $\lambda_i$  ranging over different non-zero eigenvalues and  $Q_i$  being the projectors on the respective eigenspaces. Since  $Q$  is positive semidefinite, we have  $\lambda_i > 0$  for all  $i$ .

We interpret transformations  $U_\pi$  as permutation matrices defined by  $(U_\pi)_{S,S'} = 1$  if  $S = \pi(S')$  and  $(U_\pi)_{S,S'} = 0$  otherwise. Then, we have

$$U_\pi Q U_\pi^\dagger = \frac{1}{n!} \sum_{\tau \in \mathfrak{S}_n} U_\pi P^\tau U_\pi^\dagger = \frac{1}{n!} \sum_{\tau \in \mathfrak{S}_n} P^{\pi\tau} = \frac{1}{n!} \sum_{\tau \in \mathfrak{S}_n} P^\tau = Q.$$

Since we also have

$$U_\pi Q U_\pi^\dagger = \sum_i \lambda_i U_\pi Q_i U_\pi^\dagger,$$

we must have  $Q_i = U_\pi Q_i U_\pi^\dagger$ . This means that  $Q_i$  is a projector to a subspace  $H_i \subseteq H_\varphi$  that is invariant under the action of  $\mathfrak{S}_n$ . If  $H_i$  is not irreducible, we can decompose it into a direct sum of irreducible subspaces

$$H_i = H_{i,1} \oplus H_{i,2} \oplus \dots \oplus H_{i,m_i}.$$

Then, we have  $Q_i = \sum_{j=1}^{m_i} Q_{i,j}$  where  $Q_{i,j}$  is a projector to  $H_{i,j}$  and  $Q = \sum_{i,j} \lambda_i Q_{i,j}$ . This means that we can decompose  $q(\hat{x}) = \sum_{i,j} \lambda_i q_{i,j}(\hat{x})$  where  $q_{i,j}(\hat{x}) = \vec{x}^T Q_{i,j} \vec{x}$  and it suffices to show the theorem for one polynomial  $q_{i,j}(\hat{x})$  instead of the whole sum  $q(\hat{x})$ .

### 2.2.3 Projector to one subspace.

Let  $H_{\varphi,\ell} \subseteq H_\varphi$  be an irreducible invariant subspace. We claim that the projection to the subspace  $H_{\varphi,\ell}$  denoted by  $\Pi_{\varphi,\ell}$  is of the following form:

**Lemma 2.**

$$\Pi_{\varphi,\ell} = c \rho_{\varphi,\ell} \text{ where } \rho_{\varphi,\ell} = \sum_{i_1, \dots, j_b} \vec{p}_{i_1, \dots, j_b} \vec{p}_{i_1, \dots, j_b}^T$$

for some constant  $c$ .

*Proof.* If we restrict to the subspace  $H_{\varphi,\ell}$ , then  $\Pi_{\varphi,\ell}$  is just the identity  $I$ .

On the right hand side,  $\rho_{\varphi,\ell}$  is mapped to itself by any  $U_\pi$  (since any  $U_\pi$  permutes the vectors  $\vec{p}_{i_1, \dots, j_b}$  in some way). Therefore, all  $U_\pi$  also map the eigenspaces of  $\rho_{\varphi,\ell}$  to themselves. This means that, if  $\rho_{\varphi,\ell}$  has an eigenspace  $V \subset H_{\varphi,\ell}$ , then  $U_\pi$  acting on  $V$  also form a representation of  $\mathfrak{S}_n$  but that would contradict  $H_{\varphi,\ell}$  being an irreducible representation. Therefore, the only eigenspace of  $\rho_{\varphi,\ell}$  is the entire  $H_{\varphi,\ell}$ . This can only happen if  $\rho_{\varphi,\ell}$  is  $cI$  for some constant  $c$ .  $\square$

### 2.2.4 Final polynomial

From the previous subsection, it follows that  $q_{i,j}(\hat{x})$  is a positive constant times

$$\sum_{i_1, \dots, j_b} (\hat{x}_{i_1} - \hat{x}_{j_1})^2 \dots (\hat{x}_{i_b} - \hat{x}_{j_b})^2 S^2(\hat{x}')$$

where  $S(\hat{x}')$  is a symmetric polynomial of degree at most  $t - b$ . Instead of the sum, we consider the expected value of  $(\hat{x}_{i_1} - \hat{x}_{j_1})^2 \dots (\hat{x}_{i_b} - \hat{x}_{j_b})^2 S^2(\hat{x}')$  when  $i_1, \dots, j_b$  are chosen randomly. (Since the sum and the expected value differ by a constant factor, this is sufficient.)

Terms  $(\hat{x}_{i_k} - \hat{x}_{j_k})^2$  are nonzero if and only if one of  $x_{i_k}$  and  $x_{j_k}$  is 1 and the other is  $-1$ . Then, for  $k = 1$ , we have

$$\Pr[\{\hat{x}_{i_1}, \hat{x}_{j_1}\} = \{-1, 1\}] = \frac{2s(n-s)}{n(n-1)},$$

since there are  $\frac{n(n-1)}{2}$  possible sets  $\{\hat{x}_{i_1}, \hat{x}_{j_1}\}$  and  $s(n-s)$  of them contain one 1 and one  $-1$ . For  $k > 1$ ,

$$\Pr[\{\hat{x}_{i_k}, \hat{x}_{j_k}\} = \{-1, 1\} | \{\hat{x}_{i_l}, \hat{x}_{j_l}\} = \{-1, 1\} \text{ for } l \in [k-1]]$$

$$= \frac{2(s-k+1)(n-s-k+1)}{(n-2k+2)(n-2k+1)},$$

since the condition  $\{\hat{x}_{i_l}, \hat{x}_{j_l}\} = \{-1, 1\}$  for  $l \in [k-1]$  means that, among the remaining variables, there are  $s-k+1$  variables  $\hat{x}_j = -1$  and  $n-s-k+1$  variables  $\hat{x}_j = 1$  and  $n-2k+2$  variables in total (and, given that, the  $k=1$  argument applies). Thus,

$$Pr[(\hat{x}_{i_1} - \hat{x}_{j_1})^2 \dots (\hat{x}_{i_b} - \hat{x}_{j_b})^2 = 1] = \frac{2^b s(s-1) \dots (s-b+1)(n-s) \dots (n-s-b+1)}{n(n-1) \dots (n-2b+1)}.$$

Since  $S$  is a symmetric polynomial, we have  $S(\hat{x}') = S'(s')$  where  $S'$  is a polynomial of one variable  $s'$ , with  $s'$  equal to the number of variables  $\hat{x}'_j = -1$ . Since there are  $b$  variables  $\hat{x}_j = -1$  that do not appear in  $\hat{x}'$ , we have  $s' = s - b$ . This means that  $S'$  can be rewritten as a polynomial in  $s$  (instead of  $s'$ ).

### 3 The algorithms

In Section 3.1 we now provide the algorithm for  $d \leq 3$  (the algorithm part of Theorems 1 and 2) which we know to be optimal for  $d=1$  with  $k+l=n$ , and for  $d=2,3$  and any  $k, l$ . Next, in Section 3.4 we present the sub-optimal algorithm that works for all  $d$ , resulting in a general upper bound on  $Q_E(\text{EXACT}_{k,l}^n)$  (the algorithm part of Theorem 3). Throughout Section 3 we will refer to  $\hat{x}_1 + \dots + \hat{x}_n$  as the unbalance of the input or simply unbalance, in other words, the unbalance increases as the difference between ones and zeroes in the input increases. When  $k+l=n$ , the condition  $\text{EXACT}_{k,n-k}^n(x) = 1$  is equivalent to requirement that the unbalance is  $\pm d$ , i.e.,  $|\hat{x}_1 + \dots + \hat{x}_n| = n - 2k = d$ . Hence we will refer to  $\text{EXACT}_{k,n-k}^n$  as testing for unbalance  $d = n - 2k$ .

#### 3.1 The algorithm for unbalance $d \leq 3$

For the upper bound, we now provide a quantum algorithm for the  $l = n - k$  case which can be extended to  $l \neq n - k$  case. Let us introduce the function  $\text{UNBALANCE}_d^n = \text{EXACT}_{\frac{n-d}{2}, \frac{n+d}{2}}^n$ . When  $l = n - k$  then  $d = n - 2k$  and so  $n$  and  $d$  have the same parity.

**Theorem 8.**

$$Q_E(\text{UNBALANCE}_d^n) \leq \begin{cases} \frac{n+d}{2}, & \text{if } d = 1, \\ \frac{n+d}{2} - 1, & \text{if } d \in \{2, 3\}. \end{cases}$$

We can compute  $\text{EXACT}_{k,l}^n$  for  $l \neq n - k$  by reducing it to  $\text{UNBALANCE}_{d'}^{n'}$ :

**Lemma 3.**

$$Q_E(\text{EXACT}_{k,l}^n) \leq Q_E\left(\text{UNBALANCE}_{l-k}^{n+\max\{n-l-k, l+k-n\}}\right)$$

*Proof.* For the  $l < n - k$  case ( $l > n - k$ , respectively) simply pad the input bits with  $n - l - k$  ones ( $l + k - n$  zeroes, resp.) and run  $\text{UNBALANCE}_d^{n+|n-l-k|}$  on the padded input. The complexity of the algorithm on the padded problem will be

$$Q_E(\text{EXACT}_{k,l}^n) \leq Q_E\left(\text{UNBALANCE}_d^{n+|n-l-k|}\right).$$

□

From Lemma 3 and Theorem 8, the upper bounds of Theorem 1 and Theorem 2 follow:

$$Q_E(\text{EXACT}_{k,l}^n) \leq \begin{cases} \max\{n-k, l\}, & \text{if } d = 1, \\ \max\{n-k, l\} - 1, & \text{if } d \in \{2, 3\}. \end{cases}$$



### 3.1.1 The structure of the algorithm

The algorithm of Theorem 8 will use two kinds of subroutines to calculate the function:

- The main routine  $\text{UNB}_d^n$  will start in a quantum state independent of the input and compute a  $\text{UNBALANCE}_d^n$  instance;
- The subroutine  $\text{UNB-R}_d^n$  will require a precomputed state in the form

$$\sum_{i \in [n]} \hat{x}_i |S\rangle + \sqrt{\gamma} \sum_{\substack{i, j \in [n] \\ i < j}} (\hat{x}_i - \hat{x}_j) |i, j\rangle. \quad (2)$$

$|S\rangle$  here alludes to fact that the amplitude of the basis state is a sum of  $\hat{x}_i$ 's.

Let us denote by  $\gamma(\text{UNB-R}_d^n)$  the constant coefficient  $\gamma$  of the algorithm  $\text{UNB-R}_d^n$ .

Let us denote by  $T(S)$  the number of queries performed by a subroutine  $S$ .

**Lemma 4** (Recursive step for  $\text{UNB-R}_d^n$ ). *If  $d < n$ ,  $n \geq 3$ , and there exists a quantum algorithm  $\text{UNB-R}_d^{n-2}$  computing the function  $\text{UNBALANCE}_d^{n-2}$  starting in an unnormalized quantum state of the form (2) on  $n-2$  inputs with  $\gamma(\text{UNB-R}_d^{n-2}) < 1$  then there exists an algorithm  $\text{UNB-R}_d^n$  using  $\text{UNB-R}_d^{n-2}$  as a subroutine, and computing  $\text{UNBALANCE}_d^n$ , starting in the state (2) where*

$$\gamma(\text{UNB-R}_d^n) = \frac{1}{(n^2 - d^2)^2} \left( n^2(n-2)^2 \frac{\gamma(\text{UNB-R}_d^{n-2})}{1 - \gamma(\text{UNB-R}_d^{n-2})} + d^4 \right) \quad (3)$$

and using one more query, i.e.,

$$T(\text{UNB-R}_d^n) = T(\text{UNB-R}_d^{n-2}) + 1.$$

The main routine  $\text{UNB}_d^n$  will also be recursive and make use of  $\text{UNB-R}_d^n$ .

**Lemma 5** (Recursive step for  $\text{UNB}_d^n$ ). *If there exists  $\text{UNB}_d^{n-2}$  and  $\text{UNB-R}_d^n$  with  $\gamma(\text{UNB-R}_d^n) \leq 1$ , then there exists  $\text{UNB}_d^n$  such that*

$$T(\text{UNB}_d^n) = 1 + \max\{T(\text{UNB}_d^{n-2}), T(\text{UNB-R}_d^n)\}.$$

Now we are ready to prove Theorem 8:

*Proof of Theorem 8.* We can draw the subroutine dependency graph like so:

$$\begin{array}{ccccccc} \text{UNB}_d^d & \leftarrow & \text{UNB}_d^{d+2} & \leftarrow & \text{UNB}_d^{d+4} & \leftarrow & \dots \leftarrow \text{UNB}_d^{d+2k} \\ & & \downarrow & & \downarrow & & \downarrow \\ \text{UNB-R}_d^d & \leftarrow & \text{UNB-R}_d^{d+2} & \leftarrow & \text{UNB-R}_d^{d+4} & \leftarrow & \dots \leftarrow \text{UNB-R}_d^{d+2k} \end{array}$$

Each subroutine performs one query and calls one of the subroutines in the dependency graph depending on the result of the measurement. Using Lemma 4 starting with an algorithm  $\text{UNB-R}_d^{d+2k_0}$  we can build chains of algorithms  $\text{UNB-R}_d^{d+2k_0}, \text{UNB-R}_d^{d+2(k_0+1)}, \dots, \text{UNB-R}_d^{d+2k}$  as long as  $\gamma(\text{UNB-R}_d^{d+2k_i}) < 1$ . Notice that we may use multiple chains to cover all  $k > 0$ . Fortunately, as we will show for  $d \in \{1, 2, 3\}$ , a single infinite chain will suffice.

Then, using Lemma 5 we can build algorithms  $\text{UNB}_d^{d+2k}$  for all  $k > 0$  if we additionally have an initial base algorithm for  $\text{UNB}_d^d$ . The query complexity of  $\text{UNB}_d^{d+2k}$  built in this way on a chain of  $\text{UNB-R}_d^{d+2k}$  starting at  $k_0 \in \{0, 1\}$  will have

$$T(\text{UNB}_d^{d+2k}) = \max\{k + T(\text{UNB}_d^d), T(\text{UNB-R}_d^{d+2k_0}) + k - k_0 + 1\}.$$

Since  $\text{UNB}_d^d$  is computing  $\text{EQUALITY}_d$ , it uses  $d-1$  queries, so we can disregard  $k + T(\text{UNB}_d^d)$ , since  $k = \frac{n-d}{2}$  and therefore  $k + T(\text{UNB}_d^d) \leq \frac{n+d}{2} - 1$ . To finish the proof we now need to show that there exists a chain of  $\text{UNB-R}_d^{d+2k}$  starting at  $k_0$  with  $\gamma(\text{UNB-R}_d^n) < 1$  and

$$T(\text{UNB-R}_d^{d+2k_0}) + k - k_0 + 1 \leq \begin{cases} n - k, & \text{if } d = 1, \\ n - k - 1, & \text{if } d \in \{2, 3\}. \end{cases}$$



- When  $d = 1$ , we will have  $k_0 = 0$  and show that  $T(\text{UNB-R}_d^d) \leq n - 2k + k_0 - 1 = d + k_0 - 1 = 0$ . Since the function  $\text{UNBALANCE}_1^1$  does not depend on input variables, there exists  $\text{UNB-R}_1^1$  with  $\gamma(\text{UNB-R}_1^1) = 0$  using 0 queries.
- When  $d = 2$  we will again have  $k_0 = 0$  and  $T(\text{UNB-R}_d^d) \leq d + k_0 - 2 = 0$ . The subroutine  $\text{UNB-R}_2^2$  is essentially required to compute  $XOR(x_1, x_2)$  starting in a non-normalized state  $(\hat{x}_1 + \hat{x}_2)|\mathcal{S}\rangle + \sqrt{\gamma} \cdot (\hat{x}_1 - \hat{x}_2)|1, 2\rangle$ . If  $\gamma = 0$  we can only measure  $|\mathcal{S}\rangle$  if  $XOR = 0$  and no queries are necessary.
- When  $d = 3$  a single infinite chain starting at  $k_0 = 0$  does not exist. It does exist starting at  $k_0 = 1$  and  $T(\text{UNB-R}_d^{d+2}) \leq d + k_0 - 2 = 2$ . We give algorithm for this as a separate lemma:

**Lemma 6.** *There exists a subroutine  $\text{UNB-R}_3^5$  with  $\gamma(\text{UNB-R}_3^5) = \frac{1}{112}$  using 2 queries.*

To show that the chains of algorithms  $\text{UNB-R}_d^{d+2k}$  obtained by repeated applications of Lemma 4 never have  $\gamma(\text{UNB-R}_d^{d+2k}) \geq 1$ , we use the recursive identity (3). It would be sufficient to show that  $\exists n_{\text{init}} \forall n \geq n_{\text{init}} : \gamma(\text{UNB-R}_d^n) \leq \frac{1}{n}$ . For  $n < n_{\text{init}}$  it can be verified through explicit computation. For this it would be sufficient to show that  $\exists n_{\text{init}} : \gamma(\text{UNB-R}_d^{n_{\text{init}}}) \leq \frac{1}{n_{\text{init}}} \wedge \forall n > n_{\text{init}} : \gamma(\text{UNB-R}_d^{n-2}) \leq \frac{1}{n-2} \rightarrow \gamma(\text{UNB-R}_d^n) \leq \frac{1}{n}$ . The implication holds whenever

$$\frac{\frac{n^2(n-2)^2}{n-3} + d^4}{(n^2 - d^2)^2} \leq \frac{1}{n},$$

or equivalently,

$$n^4 + (-2d^2 - 4)n^3 + (6d^2 - d^4)n^2 + 4d^4n - 3d^4 \geq 0.$$

When  $d = 1$  the inequality holds for  $n \geq 5$ . We can then numerically verify that  $\gamma(\text{UNB-R}_1^5) \approx 0.008 \leq \frac{1}{5}$ . When  $d = 2$  the inequality holds onwards from  $n \geq 12$ . For our base case  $\gamma(\text{UNB-R}_2^{12}) \approx 0.039 \leq \frac{1}{12}$ . When  $d = 3$  the inequality holds onwards from  $n \geq 23$ . For our chain  $\gamma(\text{UNB-R}_3^{23}) \approx 0.030 \leq \frac{1}{23}$ .  $\square$

### 3.2 Proof of Lemma 4

*Proof.* Our algorithm will utilize the following two unitaries and their inverses:

- $R_\alpha$  is a unitary transformation over a 3-dimensional Hilbert space with basis vectors  $|0\rangle, |\mathcal{L}\rangle$ , and  $|\mathcal{R}\rangle$ . It is a unitary completion of the following transformation:

$$R_\alpha |0\rangle = \sin \alpha |\mathcal{L}\rangle + \cos \alpha |\mathcal{R}\rangle.$$

- $U_n$  is a unitary transformation over a Hilbert space of dimension  $n + \binom{n}{2} + 1$  with basis vectors  $\{|1\rangle, |2\rangle, \dots, |n\rangle, |\mathcal{S}\rangle, |1, 2\rangle, |1, 3\rangle, \dots, |n-1, n\rangle\}$ . It is a unitary completion of the following transformation:

$$U_n |i\rangle = \frac{1}{\sqrt{n}} \left( |\mathcal{S}\rangle - \sum_{j=1}^{i-1} |j, i\rangle + \sum_{j=i+1}^n |i, j\rangle \right). \quad (4)$$

Note that on a superposition of input vectors  $U_n$  acts as:

$$U_n \sum_{i \in [n]} \alpha_i |i\rangle = \frac{1}{\sqrt{n}} \left( \sum_{i \in [n]} \alpha_i |\mathcal{S}\rangle + \sum_{\substack{i, j \in [n] \\ i < j}} (\alpha_i - \alpha_j) |i, j\rangle \right).$$

The subspace  $\{|1\rangle, \dots, |n\rangle\}$  can be regarded as the input subspace of  $U_n$  and the orthogonal subspace  $\{|\mathcal{S}\rangle, |1, 2\rangle, |1, 3\rangle, \dots, |n-1, n\rangle\}$  — as the output subspace. We will call  $|\mathcal{S}\rangle$  the sum output state and  $\{|1, 2\rangle, |1, 3\rangle, \dots, |n-1, n\rangle\}$  the difference output states. In the description of the algorithm we will specify which basis states are designated as input and output states for each  $R_\alpha$  and  $U_n$ .

We will track the state of the algorithm  $\text{UNB-R}_d^n$  throughout the recursive step. Additionally, we will introduce some real constants and specify the constraints on them induced by the algorithm. Let  $\gamma = \gamma(\text{UNB-R}_d^n)$  and  $\gamma' = \gamma(\text{UNB-R}_d^{n-2})$ . The algorithm starts in the state:

$$\sum_{i \in [n]} \hat{x}_i |\mathcal{S}\rangle + \sqrt{\gamma} \sum_{\substack{i, j \in [n] \\ i < j}} (\hat{x}_i - \hat{x}_j) |i, j\rangle.$$

We now apply  $R_\alpha$  to each of the  $|i, j\rangle$  and obtain

$$\sum_{i \in [n]} \hat{x}_i |\mathcal{S}\rangle + c_1 \sum_{\substack{i, j \in [n] \\ i < j}} (\hat{x}_i - \hat{x}_j) |i, j, \mathcal{L}\rangle + c_2 \sum_{\substack{i, j \in [n] \\ i < j}} (\hat{x}_i - \hat{x}_j) |i, j, \mathcal{R}\rangle,$$

with

$$c_1^2 + c_2^2 = \gamma. \quad (\text{C1})$$

Let us apply  $U_n^{-1}$  to the  $|\mathcal{S}\rangle$  and  $|i, j, \mathcal{L}\rangle$  parts of the state with  $|\mathcal{S}\rangle$  and  $|i, j, \mathcal{L}\rangle$  serving as the input states of  $U_n^{-1}$ . The output states for  $U_n$  are  $\{|l\rangle | l \in [n]\}$ . For each state  $|i, j, \mathcal{R}\rangle$  we perform  $U_{n-2}^{-1}$  with  $|i, j, \mathcal{R}\rangle$  serving as the sum input state of  $U_{n-2}^{-1}$  and  $|i, j, \mathcal{R}, u, v\rangle$  being some auxiliary input states with 0 amplitudes corresponding to difference input states  $\{|u, v\rangle | \{u, v\} \subseteq [n] \setminus \{i, j\}\}$ . The output states for  $|i, j\rangle$  are  $|i, j, l\rangle, l \in [n] \setminus \{i, j\}$ . We obtain:

$$c_3 \sum_{l \in [n]} \left( \sum_{i \in [n]} \hat{x}_i - c_4 \hat{x}_l \right) |l\rangle + c_5 \sum_{\substack{i, j \in [n] \\ l \in [n] \setminus \{i, j\} \\ i < j}} (\hat{x}_i - \hat{x}_j) |i, j, l\rangle.$$

It is easier to verify this statement by working backwards — pretending that we apply  $U_n$  and  $U_{n-2}$ , respectively, to the state above. Unlike their inverses, we know how to apply  $U_n$  and  $U_{n-2}$ . Combined with the following constraints, the above can be verified.

$$c_3 \cdot n - c_3 \cdot c_4 = \sqrt{n}, \quad c_3 \cdot c_4 = -c_1 \sqrt{n}, \quad c_2 = c_5 \sqrt{n-2}. \quad (\text{C2-C4})$$

The constraints can be obtained by considering the coefficient of the terms before and after the transformation. For example, the first constraint (C2) is the coefficient in front of  $\sum_{i \in [n]} \hat{x}_i$  before the transformation. If we run the algorithm backwards, the coefficient only depends on  $c_3$  and  $c_4$  from the states  $c_3 \sum_{l \in [n]} \left( \sum_{i \in [n]} \hat{x}_i - c_4 \hat{x}_l \right) |l\rangle$ .

Next, we query the variable as specified by the last number in the register, getting

$$c_3 \sum_{l \in [n]} \hat{x}_l \left( \sum_{i \in [n]} \hat{x}_i - c_4 \hat{x}_l \right) |l\rangle + c_5 \sum_{\substack{i, j \in [n] \\ l \in [n] \setminus \{i, j\} \\ i < j}} \hat{x}_l (\hat{x}_i - \hat{x}_j) |i, j, l\rangle.$$

Next, we apply  $U_n$  to the  $|l\rangle$  states as input states and using  $|\mathcal{S}\rangle$  and  $|i, j, \mathcal{L}\rangle$  as the output states. Next, for each pair  $\{i, j\}$ , we apply  $U_{n-2}$  to the group of states  $\{|i, j, l\rangle | l \in [n] \setminus \{i, j\}\}$ , thinking of those as input states and  $|i, j, \mathcal{R}\rangle$  playing the role of the sum output state  $|\mathcal{S}\rangle$  and  $|i, j, u, v\rangle$  having the role of difference output states. We obtain

$$\begin{aligned} & c_6 \left( \sum_{i, j \in [n]} \hat{x}_i \hat{x}_j - c_7 \right) |\mathcal{S}\rangle + c_8 \sum_{\substack{i, j \in [n] \\ l \in [n] \setminus \{i, j\} \\ i < j}} (\hat{x}_i - \hat{x}_j) \hat{x}_l |i, j, \mathcal{L}\rangle \\ & + c_9 \sum_{\substack{i, j \in [n] \\ l \in [n] \setminus \{i, j\} \\ i < j}} (\hat{x}_i - \hat{x}_j) \hat{x}_l |i, j, \mathcal{R}\rangle + c_{10} \sum_{\substack{i, j \in [n] \\ u, v \in [n] \setminus \{i, j\} \\ i < j \\ u < v}} (\hat{x}_i - \hat{x}_j) (\hat{x}_u - \hat{x}_v) |i, j, u, v\rangle, \end{aligned}$$

when the following constraints hold:

$$\begin{aligned} c_3 &= c_6 \sqrt{n}, & c_5 &= c_9 \sqrt{n-2}, & c_3 \cdot c_4 \cdot n &= c_6 \cdot c_7 \sqrt{n}, \\ c_3 &= c_8 \sqrt{n}, & c_5 &= c_{10} \sqrt{n-2}. \end{aligned} \quad (\text{C5-C9})$$

To finish up the unitary transformations of the recursive step we now perform  $R_\alpha^{-1}$  on the pairs of states  $|i, j, \mathcal{L}\rangle$  and  $|i, j, \mathcal{R}\rangle$  states, turning them to  $|i, j\rangle$ , and giving us

$$c_6 \left( \sum_{i,j \in [n]} \hat{x}_i \hat{x}_j - c_7 \right) |\mathcal{S}\rangle + \\ + c_{11} \sum_{\substack{i,j \in [n] \\ i < j}} (\hat{x}_i - \hat{x}_j) |i, j\rangle \left( \sum_{\substack{l \in [n] \\ l \notin \{i,j\}}} \hat{x}_l |\mathcal{S}\rangle + \sqrt{\gamma'} \sum_{\substack{u,v \in [n] \\ u,v \notin \{i,j\} \\ u < v}} (\hat{x}_u - \hat{x}_v) |u, v\rangle \right)$$

Again, this is true if the constraints obey

$$c_8^2 + c_9^2 = c_{11}^2, \quad c_{10} = c_{11} \cdot \sqrt{\gamma'}. \quad (\text{C10-C11})$$

Finally, we measure whether the state is in subspace  $\{|\mathcal{S}\rangle\}$ . We can set the constant  $c_7$  so that whenever  $\text{UNBALANCE}_d^n(x) = 1$  or equivalently  $\hat{x}_1 + \dots + \hat{x}_n = \pm d$  the amplitude of  $|\mathcal{S}\rangle$  is zero:

$$c_7 = d^2. \quad (\text{C12})$$

If on the other hand the state is not in subspace  $|\mathcal{S}\rangle$ , we end up measuring  $|i, j\rangle$  in the first register. Without loss of generality we may assume that the result is  $\{n-1, n\}$ . Thus we have learned that  $\{\hat{x}_{n-1}, \hat{x}_n\} = \{-1, 1\}$  is a balanced pair that can be removed from consideration. Furthermore, we ended up in a useful (unnormalized) state

$$\sum_{i \in [n-2]} \hat{x}_i |\mathcal{S}\rangle + \sqrt{\gamma'} \sum_{\substack{i,j \in [n-2] \\ i < j}} (\hat{x}_i - \hat{x}_j) |i, j\rangle.$$

Therefore, we can call  $\text{UNB-R}_d^{n-2}$  recursively, since

$$\text{EXACT}_{k,n-k}^n(x_1, \dots, x_n) = \text{EXACT}_{k-1,n-k-1}^{n-2}(x_1, \dots, x_{n-2}).$$

When we solve for  $\gamma$  in terms of  $n$ ,  $d$  and  $\gamma'$ , there is only one solution up to the signs of some constants  $c_i$ . The solution is specified in the statement of Lemma 4.  $\square$

### 3.3 Proof of Lemma 5

*Proof.* We start in state  $\sum_{i \in [n]} \frac{1}{\sqrt{n}} |i\rangle$ , perform the query to get  $\sum_{i \in [n]} \frac{\hat{x}_i}{\sqrt{n}} |i\rangle$  and apply  $U_n$  from Lemma 4 obtaining

$$\frac{1}{n} \left( \sum_{i \in [n]} \hat{x}_i |\mathcal{S}\rangle + \sum_{\substack{i,j \in [n] \\ i < j}} (\hat{x}_i - \hat{x}_j) |i, j\rangle \right).$$

Let  $\gamma = \gamma(\text{UNB-R}_d^n)$ . Next, we apply  $R_\gamma$  on the second part of the state, obtaining

$$\frac{1}{n} \left( \sum_{i \in [n]} \hat{x}_i |\mathcal{S}\rangle + \sqrt{\gamma} \sum_{\substack{i,j \in [n] \\ i < j}} (\hat{x}_i - \hat{x}_j) |i, j, \mathcal{L}\rangle \right) + \frac{1}{n} \sqrt{1-\gamma} \sum_{\substack{i,j \in [n] \\ i < j}} (\hat{x}_i - \hat{x}_j) |i, j, \mathcal{R}\rangle.$$

Finally we measure completely the subspace labeled with  $\mathcal{R}$ . If we obtain  $|i, j, \mathcal{R}\rangle$  we learn that  $\hat{x}_i \neq \hat{x}_j$  and thus have reduced our problem to  $\text{UNBALANCE}_d^{n-2}$  on the remaining variables which we can compute using  $\text{UNB}_{d,n-2}$ . If we obtain the orthogonal subspace, we end up in non-normalized state

$$\sum_{i \in [n]} \hat{x}_i |\mathcal{S}\rangle + \sqrt{\gamma} \sum_{\substack{i,j \in [n] \\ i < j}} (\hat{x}_i - \hat{x}_j) |i, j, \mathcal{L}\rangle,$$

which we pass to  $\text{UNB-R}_d^n$ .  $\square$

### 3.4 The general upper bound

We now present a general upper bound to  $Q(\text{EXACT}_{k,l}^n)$ . The algorithms we present are worse (by at most 2 queries) than the one in Section 3 when  $l - k = d \leq 3$ . However, they work for any  $k, l$  and thus also for any  $d$ .

First, for the special case  $k + l = n$ , we claim

**Theorem 9.**

$$Q_E(\text{EXACT}_{k,n-k}^n) \leq n - k + 1.$$

The algorithm we use to prove Theorem 9 in Section 3.4.1 involves iteratively applying a unitary, a single query, another unitary, and a measurement. On one hand, the measurement can identify a balanced pair, so we can reduce the problem size. On the other hand, it can either rule out the case  $\sum_i x_i = k$ , or the case  $\sum_i x_i = n - k$ . We then continue by solving  $\text{EXACT}_{n-k}^n$  or  $\text{EXACT}_k^n$ , respectively. The first option is favorable, as it quickly decreases the size of the remaining problem. The worst case is using the first query just to decide whether we need to be solving  $\text{EXACT}_{n-k}^n$  (or the other case). This takes further  $n - k$  queries, so the overall number of queries is bounded from above by  $1 + n - k$ .

One might wonder if this algorithm behaves any better than simply first looking at  $\text{EXACT}_k^n$  and if the answer is no, continuing with solving  $\text{EXACT}_{n-k}^n$ . It turns out that the naïve approach is not very efficient, because the algorithm for  $\text{EXACT}_k^n$  involves padding the input with extra bits. Imagine  $k = \frac{n}{3}$  and  $l = \frac{2n}{3}$ . To solve  $\text{EXACT}_k^n$ , we would need to pad the input with  $\frac{n}{3}$  bits and test for  $\text{EXACT}_{\frac{4n}{3}}^{\frac{4n}{3}}$ . We could be unlucky that  $\frac{n}{3}$  queries would just give us unbalanced pairs, always with one useless bit from the padding. After finally learning that there are *not* exactly  $k$  ones, we could scratch the  $\frac{n}{3}$  newly identified 0's in the original instance, but we would still have to continue with  $\text{EXACT}_{\frac{2n}{3}}^{\frac{2n}{3}} = \text{EQUALITY}_{\frac{2n}{3}}$ . This could require another  $\frac{2n}{3}$  queries. All in all, in the worst case we would need  $n$  queries. However, the algorithm from Section 3.4.1 uses at most  $n - k + 1$  queries, which translates to  $\frac{2n}{3} + 1$ , which is much better.

Second, for the general case  $k + l \neq n$ , we claim

**Theorem 10.**

$$Q_E(\text{EXACT}_{k,l}^n) \leq \max\{n - k, l\} + 1.$$

*Proof.* From Theorem 9 and Lemma 3:

$$Q_E(\text{EXACT}_{k,l}^n) \leq \frac{n + \max\{n - l - k, l + k - n\} + l - k}{2} + 1 = \max\{n - k, l\} + 1.$$

□

#### 3.4.1 The proof of Theorem 9: an algorithm for unbalance $\pm d$ :

In this Section we prove Theorem 9, presenting an algorithm for the problem  $\text{EXACT}_{k,n-k}^n$  that requires  $n - k + 1$  queries.

Our goal is to find an algorithm deciding whether the number of 1's in the function values is  $k$  or  $n - k$ . Equivalently, this problem can be also called  $\text{UNBALANCE}_d^n$  with  $d = l - k = n - 2k$ : does the input  $x$  have “unbalance”  $\sum_i \hat{x}_i = \pm d$  or not? This will make it easy to compare with the results of the algorithms in Section 3 for  $d = 1, 2, 3$ .

We start our algorithm with two registers prepared in the unnormalized state

$$\left( \frac{d}{n} |0\rangle + |1\rangle \right) |\mathcal{S}\rangle,$$

with  $d$  the unbalance we test for. Conditioned on the first register being  $|1\rangle$ , we transform the second register to a uniform superposition of states  $\frac{1}{\sqrt{n}} \sum_{i=1}^n |i\rangle$ . We then query the oracle. This gives us

$$\frac{d}{n} |0\rangle |\mathcal{S}\rangle + \frac{1}{\sqrt{n}} |1\rangle \sum_i \hat{x}_i |i\rangle,$$

Controlled by the first register, we apply the operation  $U_n$  from (4) to the second register (this is where another factor of  $\frac{1}{\sqrt{n}}$  comes from), producing

$$\frac{d}{n} |0\rangle |\mathcal{S}\rangle + \frac{1}{n} |1\rangle \sum_i \hat{x}_i |\mathcal{S}\rangle + \frac{1}{n} |1\rangle \sum_{i,j \in [n], i < j} (\hat{x}_i - \hat{x}_j) |i, j\rangle.$$

As we are looking at unnormalized states, we can now omit the common prefactor  $\frac{1}{n}$ . Finally, we apply a Hadamard<sup>1</sup> to the first (ancilla) register and get the unnormalized state

$$\left( \left( d + \sum_i \hat{x}_i \right) |0\rangle + \left( d - \sum_i \hat{x}_i \right) |1\rangle \right) |\mathcal{S}\rangle + (|0\rangle - |1\rangle) \sum_{i,j \in [n], i < j} (\hat{x}_i - \hat{x}_j) |i, j\rangle.$$

Finally, we measure the second register. Whenever we get a pair  $|i, j\rangle$ , we know that it is an unbalanced one, with  $\hat{x}_i = -\hat{x}_j$ . We can get rid of it, and continue solving a smaller problem with  $n' = n - 2$ . On the other hand, if we get  $|\mathcal{S}\rangle$  in the second register, we need to look at the ancilla (first) register as well. If the ancilla is  $|0\rangle$ , we learn that the overall unbalance is not  $-d$ . On the other hand, if the ancilla is  $|1\rangle$ , we learn that the overall unbalance is not  $d$ . Thus, by using a single query, our problem changes from  $\text{UNBALANCE}_d^n$  to  $\text{EXACT}_k^n$  or  $\text{EXACT}_{n-k}^n$ . Switching to the optimal algorithm for  $\text{EXACT}_k^n$ , this reduced problem can be solved in  $\leq n - k$ , i.e.  $\leq \frac{n+d}{2}$  queries.

Therefore, by iterating the above steps, we reduce the problem size by 2 several times, and then at some point reduce the problem to  $\text{EXACT}_{k'}^{n'}$  or  $\text{EXACT}_{n'-k'}^{n'}$ . The worst option in terms of the number of queries is when we never reduce the problem size, and use the very first query just to rule out one of the options  $d$  or  $-d$  for the unbalance. We then end up having to solve  $\text{EXACT}_k^n$  or  $\text{EXACT}_{n-k}^n$ , which both can use another  $n - k$  queries. Altogether, we require

$$Q_E(\text{EXACT}_{k,n-k}^n) \leq n - k + 1$$

queries. This concludes the proof of Theorem 9.

For comparison with the algorithms in Section 3, we can formulate the result in terms of the unbalance  $d$ . Recalling  $d = n - 2k$ , this algorithm finds the answer using

$$n - k + 1 = \frac{n + d}{2} + 1.$$

queries. Recall that we have  $l = n - k$  here. For  $d = 1$ , this gives  $\frac{n+3}{2}$  queries, i.e. one extra query in comparison to Theorem 8. For  $d = 2$ , this algorithm needs  $\frac{n+4}{2}$  queries, i.e. two more queries over Theorem 8. For  $d = 3$ , this algorithm needs  $\frac{n+5}{2}$  queries, i.e. again two extra queries above Theorem 8. Thus, this algorithm is not optimal for these cases and provides just an upper bound. Nevertheless, it works for general  $d$ , and thus for general  $k$ . Furthermore, by padding the input, we can get a fully general algorithm (for  $k$  and  $l$  not tied by  $l = n - k$ ) as described in the proof of Theorem 10.

## 4 Conclusion

We have shown that the exact quantum query complexity for  $\text{EXACT}_{k,l}^n$  is

$$Q_E(\text{EXACT}_{k,l}^n) = \begin{cases} \max\{n - k, l\}, & \text{if } d = 1 \text{ and } l = n - k, \\ \max\{n - k, l\} - 1, & \text{if } d \in \{2, 3\}. \end{cases}$$

where  $d = l - k$ . When  $d = 2$  and  $l = n - k$ , this provides another example of a symmetric function with  $D(f) = 2Q_E(f)$  which is the largest known gap between  $D(f)$  and  $Q_E(f)$  for symmetric functions  $f$ . To show that  $Q_E(\text{EXACT}_{k,k+1}^{2k+1}) > k$  we use an approach based on representation theory. We do not know if this lower bound method is sufficient to prove  $Q_E(f) \geq \frac{n}{2}$  for all symmetric  $f$ . In particular, it seems difficult to apply it for the symmetric function  $\text{SYM}_a$  has, for example,  $a = 0^5 1^5 0^5 1^5 0^5$ .

<sup>1</sup>Observe that the Hadamard operation is equal to  $U_2$  (4) up to up to relabeling of the basis states.

We also give a general algorithm and a lower bound, for all  $l, k$ , showing that:

$$\max\{n - k, l\} - 1 \leq Q_E(\text{EXACT}_{k,l}^n) \leq \max\{n - k, l\} + 1.$$

Previously known quantum algorithms for symmetric functions (e.g., the well known algorithm for PARITY and the algorithms for  $\text{EXACT}_k^n$  [3]) typically measure the quantum state after each query. In contrast, our algorithm for  $d \in \{1, 2, 3\}$  does not have this structure. Moreover, our numerical simulations suggest that there is no algorithm for  $\text{EXACT}_{k,l}^n$  that uses an optimal number of queries and measures the state completely after each query. We think that it is an interesting problem to study the power of quantum algorithms with the restriction that after each query the state must be measured completely and the limits of what can be achieved with such algorithms.

## Acknowledgements

This research was supported by the ERC Advanced Grant MQC, Latvian State Research Programme NexIT Project No. 1, EU FP7 project QALGO, the People Programme (Marie Curie Actions) EU's 7th Framework Programme under REA grant agreement No. 609427, Slovak Academy of Sciences, and the Slovak Research and Development Agency grant APVV-14-0878 QETWORK. We also thank Bujiao Wu (wubujiao@ict.ac.cn) for spotting the inaccuracies in the proofs of Lemmas 4 and 6.

## References

- [1] Ambainis, A.: Superlinear advantage for exact quantum algorithms. In: Proceedings of the 45th Annual ACM Symposium on Theory of Computing. pp. 891–900. ACM (2013), <http://dl.acm.org/citation.cfm?id=2488721>, arXiv preprint: <http://arxiv.org/abs/1211.0721>
- [2] Ambainis, A., Balodis, K., Belovs, A., Lee, T., Santha, M., Smotrovs, J.: Separations in query complexity based on pointer functions. Electronic Colloquium on Computational Complexity (ECCC) 22, 98 (2015), <http://eccc.hpi-web.de/report/2015/098>
- [3] Ambainis, A., Iraids, J., Smotrovs, J.: Exact quantum query complexity of exact and threshold. In: TQC. pp. 263–269 (2013), arXiv preprint: <http://arxiv.org/abs/1302.1235>
- [4] Beals, R., Buhrman, H., Cleve, R., Mosca, M., de Wolf, R.: Quantum lower bounds by polynomials. In: Proceedings of the 39th Annual Symposium on Foundations of Computer Science. pp. 352–. FOCS '98, IEEE Computer Society, Washington, DC, USA (1998), <http://dl.acm.org/citation.cfm?id=795664.796425>
- [5] Belovs, A.: Quantum algorithms for learning symmetric juntas via the adversary bound. Comput. Complex. 24(2), 255–293 (Jun 2015), <http://dx.doi.org/10.1007/s00037-015-0099-2>, arXiv preprint: <http://arxiv.org/abs/1311.6777>
- [6] Brassard, G., Høyer, P.: An exact quantum polynomial-time algorithm for Simon's problem. In: Theory of Computing and Systems, 1997., Proceedings of the Fifth Israeli Symposium on. pp. 12–23 (Jun 1997)
- [7] Cleve, R., Ekert, A., Macchiavello, C., Mosca, M.: Quantum algorithms revisited. Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences 454(1969), 339–354 (1998), <http://rspa.royalsocietypublishing.org/content/454/1969/339.short>
- [8] Deutsch, D., Jozsa, R.: Rapid solution of problems by quantum computation. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 439(1907), 553–558 (1992), <http://rspa.royalsocietypublishing.org/content/439/1907/553>
- [9] Farhi, E., Goldstone, J., Gutmann, S., Sipser, M.: Limit on the speed of quantum computation in determining parity. Phys. Rev. Lett. 81, 5442–5444 (Dec 1998), <http://link.aps.org/doi/10.1103/PhysRevLett.81.5442>, arXiv preprint: <http://arxiv.org/abs/quant-ph/9802045>

- [10] von zur Gathen, J., Roche, J.R.: Polynomials with two values. *Combinatorica* 17(3), 345–362 (1997), <http://link.springer.com/article/10.1007/BF01215917>
- [11] Lee, T., Prakash, A., de Wolf, R., Yuen, H.: On the Sum-of-Squares Degree of Symmetric Quadratic Functions. In: Raz, R. (ed.) 31st Conference on Computational Complexity (CCC 2016). *Leibniz International Proceedings in Informatics (LIPIcs)*, vol. 50, pp. 17:1–17:31. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany (2016), <http://drops.dagstuhl.de/opus/volltexte/2016/5838>
- [12] Midrijānis, G.: Exact quantum query complexity for total boolean functions. *arXiv preprint quant-ph/0403168* (2004), <http://arxiv.org/abs/quant-ph/0403168>
- [13] Montanaro, A., Jozsa, R., Mitchison, G.: On exact quantum query complexity. *Algorithmica* 71(4), 775–796 (2015), <http://dx.doi.org/10.1007/s00453-013-9826-8>, *arXiv preprint*: <http://arxiv.org/abs/1111.0475>
- [14] Qiu, D., Zheng, S.: Characterizations of symmetrically partial boolean functions with exact quantum query complexity. *arXiv preprint abs/1603.06505* (2016), <http://arxiv.org/abs/1603.06505>

## A Proof of Lemma 6

*Proof.* The algorithm is similar to the subroutine  $\text{UNB-R}_d^n$ . The goal is to construct an amplitude that is a symmetric polynomial of degree 3 of the form  $\sum_{i < j < k} \hat{x}_i \hat{x}_j \hat{x}_k + c \sum_i \hat{x}_i$  that is 0 when  $|x| \in \{1, 4\}$ . Start with

$$\sum_{i \in [n]} \hat{x}_i |S\rangle + c_1 \sum_{\substack{i, j \in [n] \\ i < j}} (\hat{x}_i - \hat{x}_j) |i, j\rangle.$$

and perform  $R_\alpha$  for suitable  $\alpha$  on  $c_1$  part of the state. Then  $U_5^{-1}$  to obtain  $c_2$  part of the state and  $U_3^{-1}$  on the remainder to obtain  $c_4$  part of the state.

$$c_2 \sum_{i \in [n]} \left( \sum_{\substack{j \in [n] \\ j \neq i}} \hat{x}_j + c_3 \hat{x}_i \right) |i\rangle |S\rangle + c_4 \sum_{k \in [n]} \sum_{\substack{i, j \in [n] \setminus \{k\} \\ i < j}} (\hat{x}_i - \hat{x}_j) |k\rangle |i, j\rangle$$

The constraints induced by these transformations are

$$(c_2 \cdot c_3 + 4 \cdot c_2) / \sqrt{5} = 1, \quad c_1^2 = (c_2 \cdot (c_3 - 1) / \sqrt{5})^2 + (3 \cdot c_4 / \sqrt{3})^2.$$

We continue with a query on variable indicated by the first register.

$$c_2 \sum_{i \in [n]} \left( \hat{x}_i \sum_{\substack{j \in [n] \\ j \neq i}} \hat{x}_j + c_3 \right) |i\rangle |S\rangle + c_4 \sum_{k \in [n]} \hat{x}_k \sum_{\substack{i, j \in [n] \setminus \{k\} \\ i < j}} (\hat{x}_i - \hat{x}_j) |k\rangle |i, j\rangle$$

Next, we perform  $U_5$  on  $c_2$  part of the state to obtain  $c_5$  and some  $c_7$  and  $U_3$  on  $c_4$  part of the state to obtain some  $c_7$  and  $c_8$ . Perform  $R_\alpha^{-1}$  for suitable  $\alpha$  to merge  $c_7$  states.

$$\begin{aligned} c_5 \left( \sum_{\substack{i, j \in [n] \\ i < j}} \hat{x}_i \hat{x}_j + c_6 \right) |S\rangle + c_7 \sum_{\substack{i, j \in [n] \\ i < j}} (\hat{x}_i - \hat{x}_j) \left( \sum_{\substack{k \in [n] \\ k \notin \{i, j\}}} \hat{x}_k \right) |i, j\rangle + \\ + c_8 \sum_{\substack{i, j, k, l \in [n] \\ i < j \\ k < l \\ \{i, j\} \cap \{k, l\} = \emptyset}} (\hat{x}_i - \hat{x}_j)(\hat{x}_k - \hat{x}_l) |i, j, k, l\rangle \end{aligned}$$

The constraints induced by these transformations are

$$\begin{aligned} 5 \cdot c_2 \cdot c_3 / \sqrt{5} &= c_5 \cdot c_6, & c_2 \cdot 2 / \sqrt{5} &= c_5, \\ c_7^2 &= c_2^2 / 5 + c_4^2 / 3, & c_4 / \sqrt{3} &= c_8. \end{aligned}$$



First we split  $|i, j\rangle$  into two parts using  $R_\alpha$ . Now we perform  $U_5^{-1}$  on  $|\mathcal{S}\rangle$  as input sum state and  $c|i, j\rangle$  for some  $c$  as difference states. The output is stored as  $|i\rangle|\mathcal{S}\rangle$  states. For each  $|i, j\rangle$  we apply  $U_3^{-1}$  to  $\sqrt{c_7 - c^2}|i, j\rangle$  as input sum state and  $|i, j, k, l\rangle$  as the difference states. Again, it is easier to verify that we obtain the following state by running the algorithm backwards.

$$c_9 \sum_i \left( \hat{x}_i \sum_{j \neq i} \hat{x}_j + c_{10} \sum_{\substack{j < k \\ j \neq i \\ k \neq i}} \hat{x}_j \hat{x}_k + c_{11} \right) |i\rangle |\mathcal{S}\rangle + \\ + c_{12} \sum_k \sum_{\substack{i < j \\ i \neq k \\ j \neq k}} (\hat{x}_i - \hat{x}_j) \left( \sum_{l \notin \{i, j, k\}} \hat{x}_l + c_{13} \hat{x}_k \right) |k\rangle |i, j\rangle$$

This step induces the constraints:

$$(2 \cdot c_9 + 3 \cdot c_9 \cdot c_{10})/\sqrt{5} = c_5, \\ c_9 \cdot c_{11} \cdot 5/\sqrt{5} = c_5 \cdot c_6, \\ c_7^2 = (c_9 \cdot (1 - c_{10})/\sqrt{5})^2 + (c_{12} \cdot (2 + c_{13})/\sqrt{3})^2, \\ c_8 = c_{12} \cdot (c_{13} - 1)/\sqrt{3}.$$

Perform a query on variable indicated by the first register.

$$c_9 \sum_i \left( \sum_{j \neq i} \hat{x}_j + c_{10} \hat{x}_i \sum_{\substack{j < k \\ j \neq i \\ k \neq i}} \hat{x}_j \hat{x}_k + c_{11} \hat{x}_i \right) |i\rangle |\mathcal{S}\rangle + \\ + c_{12} \sum_k \sum_{\substack{i < j \\ i \neq k \\ j \neq k}} (\hat{x}_i - \hat{x}_j) \left( \hat{x}_k \sum_{l \notin \{i, j, k\}} \hat{x}_l + c_{13} \right) |k\rangle |i, j\rangle$$

Apply  $U_5$  to the  $|i\rangle|\mathcal{S}\rangle$  states as input states and  $|\mathcal{S}\rangle$  as the output sum state and part of  $|i, j\rangle$  as output difference states. For each  $i, j$  run  $U_3$  on  $|k\rangle|i, j\rangle$  states as input and  $|i, j\rangle$  as the output sum state and  $|i, j, k, l\rangle$  as the difference states. Notice, that we obtain  $|i, j\rangle$  states from two different sources — as sum states from  $|k\rangle|i, j\rangle$  and difference states from  $|i\rangle|\mathcal{S}\rangle$ . We then merge them using  $R_\alpha^{-1}$  for suitable  $\alpha$ , getting

$$c_{14} \left( \sum_{i < j < k} \hat{x}_i \hat{x}_j \hat{x}_k + c_{15} \sum_i \hat{x}_i \right) |\mathcal{S}\rangle + \\ + c_{16} \sum_{i < j} (\hat{x}_i - \hat{x}_j) \left( \sum_{\substack{k < l \\ \{i, j\} \cap \{k, l\} = \emptyset}} \hat{x}_k \hat{x}_l + c_{17} \right) |i, j\rangle + \\ + c_{18} \sum_{\substack{i, j, k, l, m \\ i < j \\ k < l \\ \{i, j\} \cap \{k, l\} = \emptyset \\ m \notin \{i, j, k, l\}}} (\hat{x}_i - \hat{x}_j)(\hat{x}_k - \hat{x}_l) \hat{x}_m |i, j, k, l\rangle.$$

This step induces the constraints:

$$c_{14} = 3 \cdot c_9 \cdot c_{10}/\sqrt{5}, \\ c_{14} \cdot c_{15} = (4 \cdot c_9 + c_9 \cdot c_{11})/\sqrt{5}, \\ c_{16}^2 = (c_{12} \cdot 2/\sqrt{3})^2 + (c_9 \cdot c_{10}/\sqrt{5})^2,$$

$$\begin{aligned}
c_{18} &= c_{12}/\sqrt{3}, \\
c_{11} - 1 &= c_{17} \cdot c_{10}, \\
3c_{13} &= 2c_{17}.
\end{aligned}$$

The amplitude of the  $c_{14}$  part is zero, when  $|x| \in \{1, 4\}$ . The amplitude of the  $c_{16}$  part is zero, when  $|x| \in \{0, 2, 3, 5\}$ . The amplitude of the  $c_{18}$  part is zero, when  $|x| \in \{0, 1, 4, 5\}$ . This induces constraints:

$$-2 + c_{15} \cdot 3 = 0, \quad -1 + c_{17} = 0.$$

There is only one solution to the system of constraints up to some  $\pm$  signs.

$$\begin{aligned}
c_1 &= 1/(4\sqrt{7}), & c_{10} &= 5, \\
c_2 &= 17/(16\sqrt{5}), & c_{11} &= 6, \\
c_3 &= 12/17, & c_{12} &= (3\sqrt{3/7})/16, \\
c_4 &= \sqrt{3/7}/16, & c_{13} &= 2/3, \\
c_5 &= 17/40, & c_{14} &= 3/8, \\
c_6 &= 30/17, & c_{15} &= 2/3, \\
c_7 &= (2\sqrt{2/7})/5, & c_{16} &= 1/(2\sqrt{7}), \\
c_8 &= 1/(16\sqrt{7}), & c_{17} &= 1, \\
c_9 &= 1/(8\sqrt{5}), & c_{18} &= 3/(16\sqrt{7}).
\end{aligned}$$

□

## B An asymptotically optimal algorithm for a class of symmetric functions

*Proof of Theorem 4.* The lower bound follows trivially from von zur Gathen's and Roche's work [10] and the polynomial method [4].

For the upper bound, we now present an algorithm. Its main idea is to use the algorithm for UNBALANCE $_d^n$  from Section 3.4.1 to successively eliminate the weights  $i$  from consideration, such that  $a_i = 1$ , or reduce the problem size by 2 bits. Once we have eliminated all possible weights  $i : a_i = 1$  except one, we switch to the algorithm for EXACT $_k^n$ . Thus the algorithm consists of two stages.

In the first stage of the algorithm we will keep track of the problem, to which we have reduced the original problem, by  $\mathbf{a} \in \{0, 1, *\}^{n+1}$ , where  $\mathbf{a}_i = *$  indicates  $|x'| \neq i$  where  $x'$  is the padded input. Padding input  $x$  with a zero corresponds to appending 0 at the end of  $\mathbf{a}$ ; padding with a one corresponds to prepending 0 at the beginning of  $\mathbf{a}$ . Let us denote this operation by  $\mathbf{a} : 0$  or  $0 : \mathbf{a}$ . A useful operation, when we learn that  $x_i \neq x_j$ , is the removal of first and last elements of  $\mathbf{a}$  and it will be denoted by  $\downarrow \mathbf{a} \downarrow$ . The algorithm starts with  $\mathbf{a} = a$ . If at any point during this computation  $\mathbf{a}$  does not contain 0 (or 1, respectively), we output 1 (0, resp.). If we never reach the point where  $\mathbf{a}$  does not contain 0 or 1, but the first stage finishes, in the end  $\mathbf{a}$  must contain exactly one 1. This is so, because the pointer to the middle  $m$  sweeps through all 1-s originally in  $a$ .

Now we proceed with the second stage of our algorithm. To decide  $\mathbf{a}$ , knowing that for exactly one weight  $i : \mathbf{a}_i = 1$ , we use an algorithm for EXACT $_{k'}^{n'}$  where  $n'$  is the size of the input at the end of the first stage and  $k'$  is the only weight with  $\mathbf{a}_{k'} = 1$ .

Let us calculate the number of queries used by the first stage.

$$\#(\text{queries used by Stage 1}) = \#(\text{we hit branch } |x| \neq a_i) + \#(\text{we hit branch } x_k \neq x_l).$$

Since the initial  $a$  contains at most  $2g(n) + 1$  ones,  $\#(\text{we hit branch } |x| \neq a_i) \leq 2g(n)$ . For the sake of brevity, let us denote  $\#(\text{we hit branch } x_k \neq x_l)$  by  $t$ . The number of queries used by second stage of the algorithm depends on  $t$ , because we have eliminated  $2t$  variables from  $x$ , padded or otherwise.

---

**Algorithm 1** Stage 1: Eliminate multiple weights  $i : a_i = 1$ 


---

```

for  $i = 1, 2, \dots, 2g(n)$  do
     $a \leftarrow 0 : a$ 
end for
 $\ell \leftarrow \frac{n}{2} + g(n)$   $\triangleright \ell$  points to the position of leftmost potential 1
for  $m = 0, \frac{1}{2}, 1, \dots, 2g(n) - \frac{1}{2}, 2g(n)$  do  $\triangleright m$  is the middle of  $a$  relative to  $\ell$ 
    while  $\exists a_i = a_j = 1 : \frac{i+j}{2} = \ell + m$  do
        Run one step of the  $\text{UNBALANCE}_{j-i}^{2\ell+2m}$  algorithm from Section 3.4.1
        if we learn  $x_k \neq x_l$  then
             $a \leftarrow a \downarrow$ 
             $\ell \leftarrow \ell - 1$ 
        else  $\triangleright$  We learn that  $|x'| \neq a_i$ 
             $a_i \leftarrow *$ 
        end if
    end while
     $a \leftarrow a : 0$ 
end for

```

---

Now, to determine the complexity of the second stage, we calculate  $n' = n + 6g(n) + 1 - 2t$  and  $\frac{n}{2} + g(n) - t \leq k' \leq \frac{n}{2} + 3g(n) - t$ . Then,

$$\#(\text{queries used by Stage 2}) = Q_E(\text{EXACT}_{k'}^{n'}) \leq \max \left\{ \frac{n}{2} + 5g(n) + 1 - t, \frac{n}{2} + 3g(n) - t \right\}.$$

The total number of queries we use is at most

$$\begin{aligned} & \#(\text{queries used by Stage 1}) + \#(\text{queries used by Stage 2}) \leq \\ & \leq 2g(n) + t + \frac{n}{2} + 5g(n) + 1 - t = \frac{n}{2} + 7g(n) + 1 = \frac{n}{2} + o(n) \sim \frac{n}{2} \end{aligned}$$

□

The algorithm from Section 3.4.1 can be extended to either exclude one of two weights for  $\text{EXACT}_{k,l}^n$  or two opposite inputs for more general  $k, l$ , as described in Section C. The only requirement for the weights  $k, l$  imposed by the algorithm is that  $k < \frac{n}{2} < l$ . Using that algorithm, the constant  $c$  in front of  $g(n)$  can be decreased from 7 to 5. The constant is relevant when  $g(n) = \epsilon n$  since it lets us construct exact quantum algorithms using less than  $n$  queries, provided  $\epsilon < \frac{1}{2c}$ .

## C Another algorithm: testing for unbalance $+u, -w$

In this Section, we develop a test for two particular, nonzero values of unbalance with opposite signs, i.e. solving the problem  $\text{EXACT}_{\frac{n-u}{2}, \frac{n+w}{2}}^n$  for  $0 < u, w \leq n$ . It can be later used quite efficiently to test for a range of unbalances, eliminating their extreme values one by one (from either end, at random), or decreasing the problem size. The test generalizes the approach of Section 3.4.

### C.1 Testing for two particular values $\{+u, -w\}$ of the unbalance.

Our goal is to determine whether for a function with exactly  $k$  ones, the unbalance  $\sum_i \hat{x}_i = -k + n - k = n - 2k$  is exactly  $+u$  or  $-w$ , with  $u, w > 0$ . We will do this by an algorithm that repeats a number of steps that either end up reducing the problem size, eliminating the option  $+u$ , or eliminating the option  $-w$ .

Let us start with an unnormalized, two-register (ancilla, data) state

$$\left( \frac{\sqrt{uw}}{n} |0\rangle + |1\rangle \right) |\mathcal{S}\rangle.$$

Controlled on the first, ancilla register being  $|1\rangle$ , we prepare a uniform superposition in the second, data register and query the data register. We obtain

$$\frac{\sqrt{uw}}{n} |0\rangle |\mathcal{S}\rangle + \frac{1}{\sqrt{n}} |1\rangle \sum_i \hat{x}_i |i\rangle$$

Next, again controlled on the ancilla being  $|1\rangle$ , we apply the unitary  $U_n$  (4), giving us

$$\frac{\sqrt{uw}}{n} |0\rangle |\mathcal{S}\rangle + \frac{1}{n} \left( \sum_i \hat{x}_i \right) |1\rangle |\mathcal{S}\rangle + \frac{1}{n} |1\rangle \sum_{i < j} (\hat{x}_i - \hat{x}_j) |i, j\rangle.$$

We look at unnormalized states, so we can drop the overall normalization factor  $\frac{1}{n}$ . Finally, conditioned on the second register being  $|\mathcal{S}\rangle$ , we rotate the first qubit using the unitary

$$Q = \frac{1}{\sqrt{u+w}} \begin{bmatrix} \sqrt{u} & -\sqrt{w} \\ \sqrt{w} & \sqrt{u} \end{bmatrix},$$

This results in the (unnormalized) state

$$\frac{1}{\sqrt{u+w}} \left( \sqrt{w} \left( u - \sum_i \hat{x}_i \right) |0\rangle + \sqrt{u} \left( w + \sum_i \hat{x}_i \right) |1\rangle \right) |\mathcal{S}\rangle + |1\rangle \sum_{i < j} (\hat{x}_i - \hat{x}_j) |i, j\rangle.$$

Finally, we perform a full measurement. Whenever we get a pair  $|i, j\rangle$  in the second register, we know that it is unbalanced, with  $\hat{x}_i = -\hat{x}_j$ . We can get rid of it, and continue solving a smaller problem with  $n' = n - 2$ , and the same possible unbalances. On the other hand, if the second register is  $|\mathcal{S}\rangle$ , the value of the first register tells us that the unbalance is either not equal to  $u$  (if we measure  $|0\rangle$ ), or not equal to  $-w$  (if we measure  $|1\rangle$ ).

## C.2 Computing $\text{SYM}_a$

We can use the algorithm described in this appendix (instead of the algorithm from Section 3.4.1) as a subroutine to the algorithm Appendix B. We also slightly modify the algorithm from Appendix B. Now instead of moving  $m$  through all of the 1-s in  $a$ , we start with  $l + m = \frac{n}{2}$ . Then call algorithm from Appendix C until either 1-s left of  $m$  are eliminated or 1-s to the right of  $m$  are eliminated. Then proceed by moving  $m$  in the direction of the remaining 1-s.

Again, the algorithm has two stages with complexities:

$$\#(\text{queries used by Stage 1}) = \#(\text{we hit branch } |x| \neq a_i) + \#(\text{we hit branch } x_k \neq x_l);$$

and if we denote  $\#(\text{we hit branch } x_k \neq x_l)$  in the Stage 1 by  $t$ , the number of variables in the input after Stage 1 by  $n' = n + 2g(n) - 2t$  and by  $k'$  the only weight for which  $\mathbf{a}_{k'} = 1$ ,  $\frac{n}{2} - g(n) - t \leq k' \leq \frac{n}{2} + 3g(n) - t$  then

$$\#(\text{queries used by Stage 2}) = Q_E(\text{EXACT}_{k'}^{n'}) \leq \frac{n}{2} + 3g(n) - t.$$

Thus the total number of queries used is  $\leq \frac{n}{2} + 5g(n)$ .